

本周周报（2013.9.9-2013.9.15）

郭方舟

本周工作

1. VAST Challenge 3 review 总结

a. 阅读了 VAST Challenge 3 的 review，主要问题总结如下：

1) 对隐含事件的探测很差

在整个分析结果中，最终只能找到 DOS 攻击，其他类型的攻击事件都没有找到。

2) 对事件的分析比较差，缺乏专业背景，使第二问的回答完全不着调

3) 对数据的利用不够

只利用了 netflow 数据中的统计信息，而忽略了 netflow 中的其他字段以及另外两个数据集。

4) 径向排布的可视化设计难以理解

多个 review 提到在提交的答案中缺乏对可视设计的解释，可视设计难以理解。

5) 缺乏丰富的交互，系统易用度低

b. 这次完成的 VAST Challenge 3 的问题非常多，对于整个过程中的问题总结如下：

1) 对本次 VAST Challenge 3 的重视度不够

这是没有做好的最主要的原因。由于身处毕业季，没能将全部精力投入到项目中来，使得整个项目的完成度不高。

2) 计算机安全方面知识匮乏

缺乏经验，应该首先找到一些计算机安全方面的论文、书籍读一读。在整个项目进行过程中，对于网络事件的分析一直处于盲目猜测的水平，缺乏基于事实的分析，在今后的项目中应避免这种情况出现。

3) 时间安排不合理

前期大量的时间消耗在了找寻数据中的规律，但是由于数据库设计不合理，导致在消耗时间的同时，又没有找到真正有价值的规律，使得后期实现、分析过程中困难重重。

4) 对大规模数据的处理经验不够

如前文所说，初始的数据库设计过于简单，导致查询速度极慢，拖延了项目进度。

c. 对于本项目未来工作的一些想法：

1) VAST Challenge 3 给出的数据类型本质上是大量的连接信息以及在每个连接节点的附加信息，这种类型的数据与社交网络数据，个人信息空间数据（比如 E-mail 数据）并没有太大的区别。未来工作有没有可能可以围绕这一整类数据展开，参考《A Model for Structure-based Comparison of Many Categories in Small-Multiple Displays》做一个可视化模型或是可视化流程。

2) 我与一个数据库实验室的同学交流了一下，这位同学说 Oracle 数据库对大规模数据的处理效果较 MySQL 数据库要好。未来我想要将现在的数据放到 Oracle 中，获得更强的实时性，以满足未来更多地交互操作。

3) 从上个月做出的体绘制结果来看，使用邻接矩阵数据来进行体绘制的结果太过稀疏。可视化结果太稀疏、空间浪费比较严重这一点在 review 中也有提到。我觉得应该想一种既能够表示出连接变化，又能够使体绘制结果变得更加紧凑的布局形式。

2. 阅读论文《A Model for Structure-based Comparison of Many Categories in Small-Multiple Displays》并做组会报告。

3. 阅读《Crafting Your Research Future》。

上个月在完成了 VAST Challenge 3 的 two-page summary。虽然只有短短两页，但是也体会到了使用英文写论文的难度。在阅读了《Crafting Your Research Future》第六章以及上个月看的《How to Write a Paper》之后，有一些体会，总结如下：

1) 写论文最重要的是能让别人看懂，表达自己对一个问题的看法以及解决这个问题思路。

写论文不是出阅读理解题，句子结构应尽量简单，多用主动式，降低读者的理解难度。

2) 论文的 abstract 非常重要。

abstract 可以传达给读者整篇论文的核心内容，相当于两个人见面时的第一印象。

3) 还记录了一些文中提到的表达方法以及惯用词。

下周工作

1. 上课。
2. 最新的 visweek 论文已经出炉，挑选几篇论文进行精读。
3. 进行 vast challenge 的改进工作。